



CHAT-PUB

Helper of welfare policy for youth

김강산
양승빈
박진호
전창민



TABLE OF CONTENTS

 **Introduction**

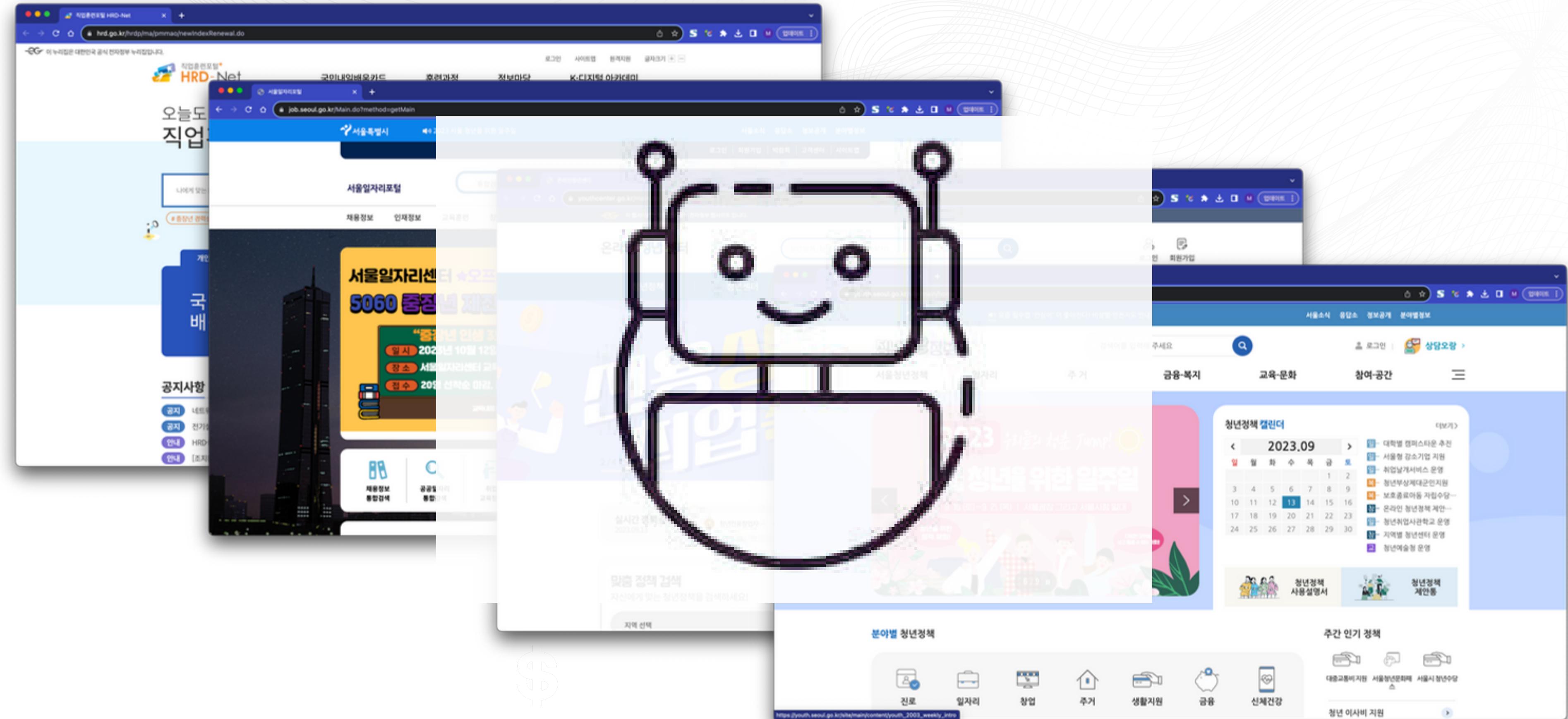
 **Implementation**

- Data Generation
- Model
- Back-End
- Front-End

 **Schedule**

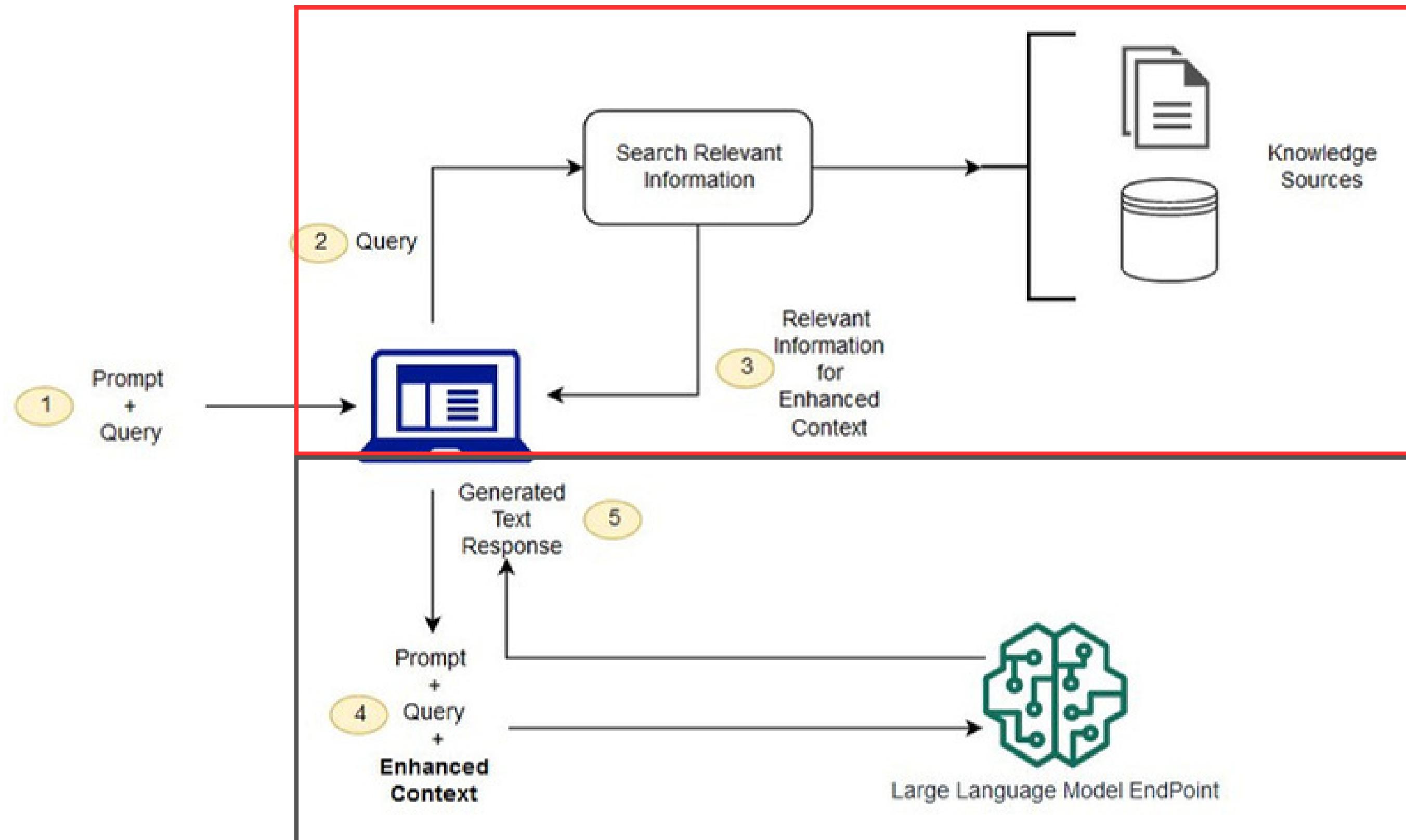


INTRODUCTION



MODEL RECAP

RETRIEVAL AUGMENTED GENERATION



[Retrieval]

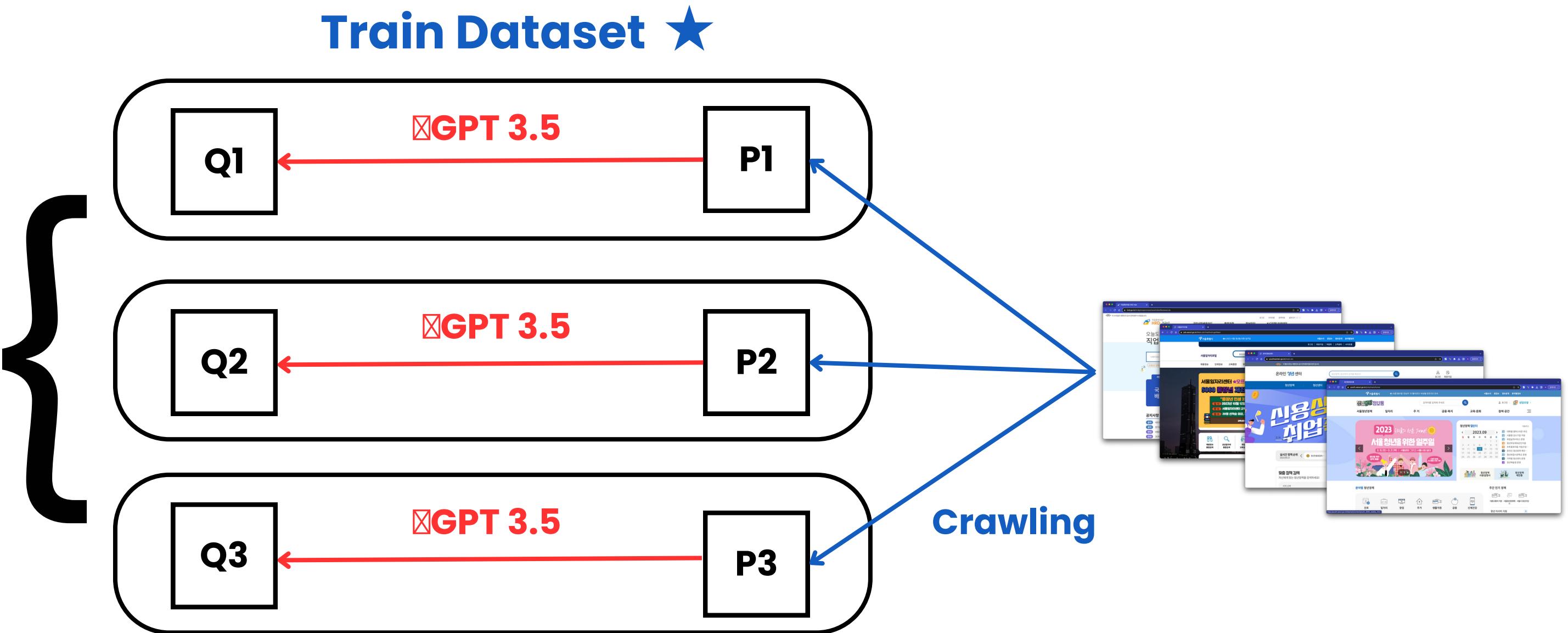
**Sentence
Transformers**

[Generation]

**Open AI API
(GPT-4)**

DATA GENERATION

*Train
sentence
transformer!!*



EXAMPLE

Q1: What's Sungkyunkwan University?
P1: Sungkyunkwan University is one of the best universities in Korea.

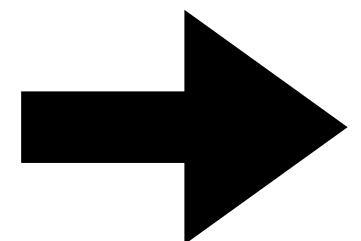
Generating Datasets with Pretrained Language Models
GPT3.5

DATA GENERATION



GPT3.5

well prompt...



```
{  
    "question": "안녕?",  
    "passage": "안녕하세요."  
,  
    {  
        "question": "오늘의 날씨는 어때?",  
        "passage": "매우 춥고 바람이 많이 불어요."  
,  
        {  
            "question": "취미가 뭐야?",  
            "passage": "책 읽기."  
        }  
    }  
}
```

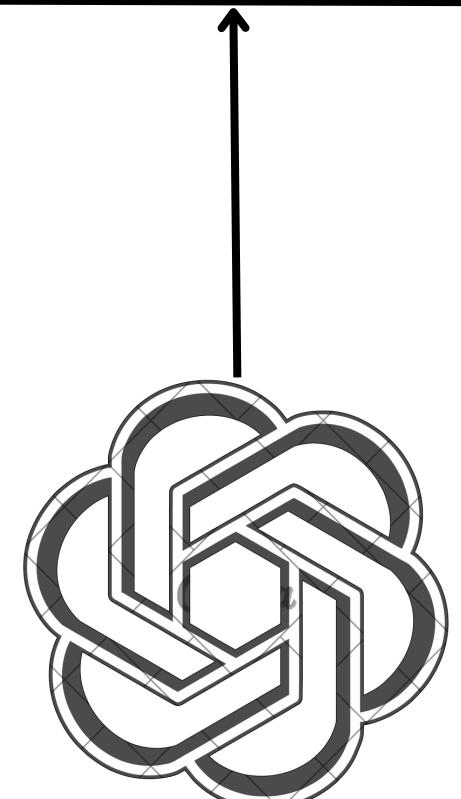
DATA GENERATION

POLICY DATA [PASSAGE]

- 제목: 미취업 청년 어학 자격시험 응시료 지원사업
- 지원내용: 미취업 청년 고른 취업기회 제공과 역량강화 비용 지원
- 신청 절차: 신청일 당시 도내 주민등록지에 신청
- 연령: 만 19세 ~ 34세
- 거주지 및 소득: 도내 거주 미취업 청년(만

[few-shot prompting]
You are an expert in question generation.
Your task is, when you receive information about Korea's youth policy,

BEST
PROMPT



OUTPUT [QUESTION : PASSAGE]

- Q1: 경기도에 거주하고 미취업한 25세 청년이 어학자격시험 응시료를 지원받을 수 있는지 알려줄 수 있을까요? P1: 제목, 연령, 거주지 및 소득
- Q1: 청년 어학시험 응시료를 지원받을 수 있는 기간은 언제부터 언제까지인가요? P1: 제목, 정책분야, 사업운영기간, 사업 신청 기간

DATA GENERATION

```
def generate_question(passage):
    messages = [
        {'role': 'system', 'content': system_prompt},
        {'role': 'user', 'content': user_prompt.format(passage)}
    ]
    response = openai.ChatCompletion.create(
        model="gpt-3.5-turbo",
        messages=messages,
        temperature=0,
        max_tokens=100,
        top_p=1.0,
        frequency_penalty=0,
        presence_penalty=0
    )
    return response['choices'][0]['message']['content']
```

[system prompt]

You are an expert in question generation.
Your task is, when you receive information
about Korea's youth policy,
generate question that can deduce the
answer from the policy information.

예시 1:

"지원내용: 미취업 청년 고른 취업기회 제공과
역량강화 비용 지원"

Example for output :

Q: 현재 미취업상태 청년들에게 해당되는 혜택
이 있는지 알려주세요....

[user prompt]]

Make one question using the policy
information given below

정책 정보: {passage}

DATA GENERATION

• 1220 DATASETS

```
{ } data1206.json  
{ } data1207.json  
{ } data1208.json  
{ } data1209.json  
{ } data1210.json  
{ } data1211.json  
{ } data1212.json  
{ } data1213.json  
{ } data1214.json  
{ } data1215.json  
{ } data1216.json  
{ } data1217.json  
{ } data1218.json  
{ } data1219.json  
{ } data1220.json
```

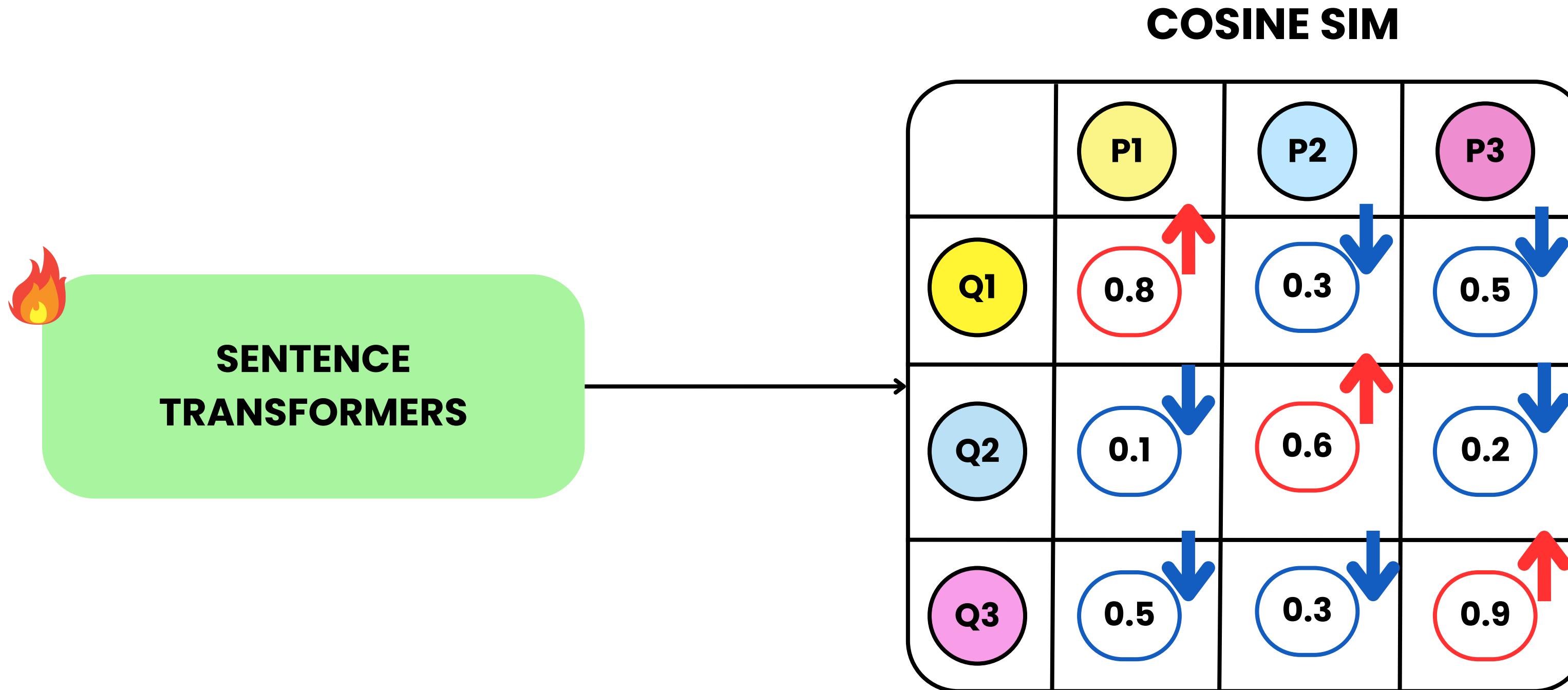
• DATA1216.JSON

```
{  
    "question": "Q: 경제적으로 어려운 장애 청년들을 위해 어떤 지원 정책이 있을까요?",  
    "passage": "지원내용: 경제적 취약 계층인 장애 청년의 자립기반 마련을 위해 2년간 월 10만 원의 생활비 지원",  
},  
{  
    "question": "Q: 화성형 장애인 누림통장을 신청하기 위해 필요한 서류는 무엇인가요?",  
    "passage": "제목: 화성형 장애인 누림통장 운영\\n신청 절차: 주소지 관할 행정복지센터 신청",  
},  
{  
    "question": "Q: 화성시에 주민등록을 두고 있는 만 19세 등록장애인의 화성형 장애인 누림통장을 신청하는 경우 연령은?",  
    "passage": "제목: 화성형 장애인 누림통장 운영\\n연령: 만 19세 ~ 19세\\n거주지 및 주민등록증",  
},  
{  
    "question": "Q: 장애인 누림통장 운영 정책은 어떤 내용으로 지원을 해주는 건가요?",  
    "passage": "제목: 화성형 장애인 누림통장 운영\\n정책 번호: R2023060213643\\n정책 내용: 경제적 취약 계층인 장애 청년의 자립기반 마련을 위한 지원 정책",  
}
```

MODEL

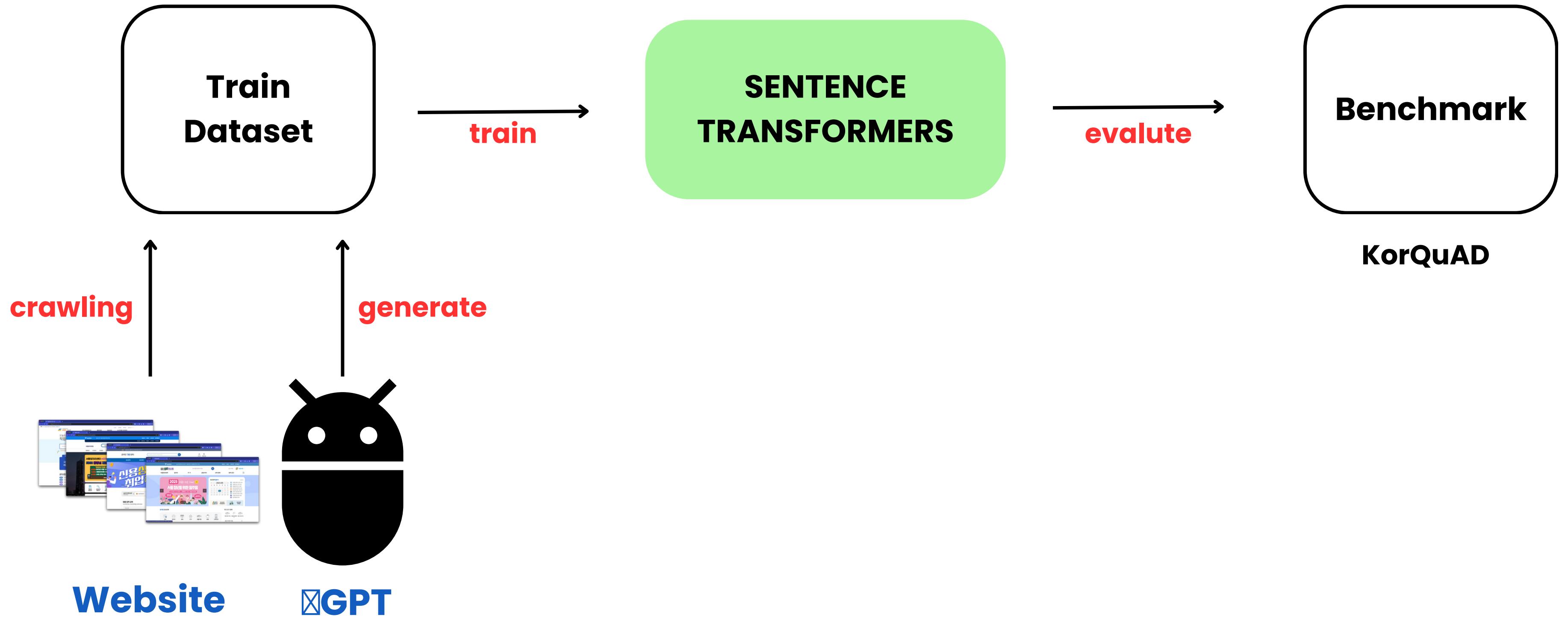
RECAP: SENTENCE TRANSFORMERS

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|}$$



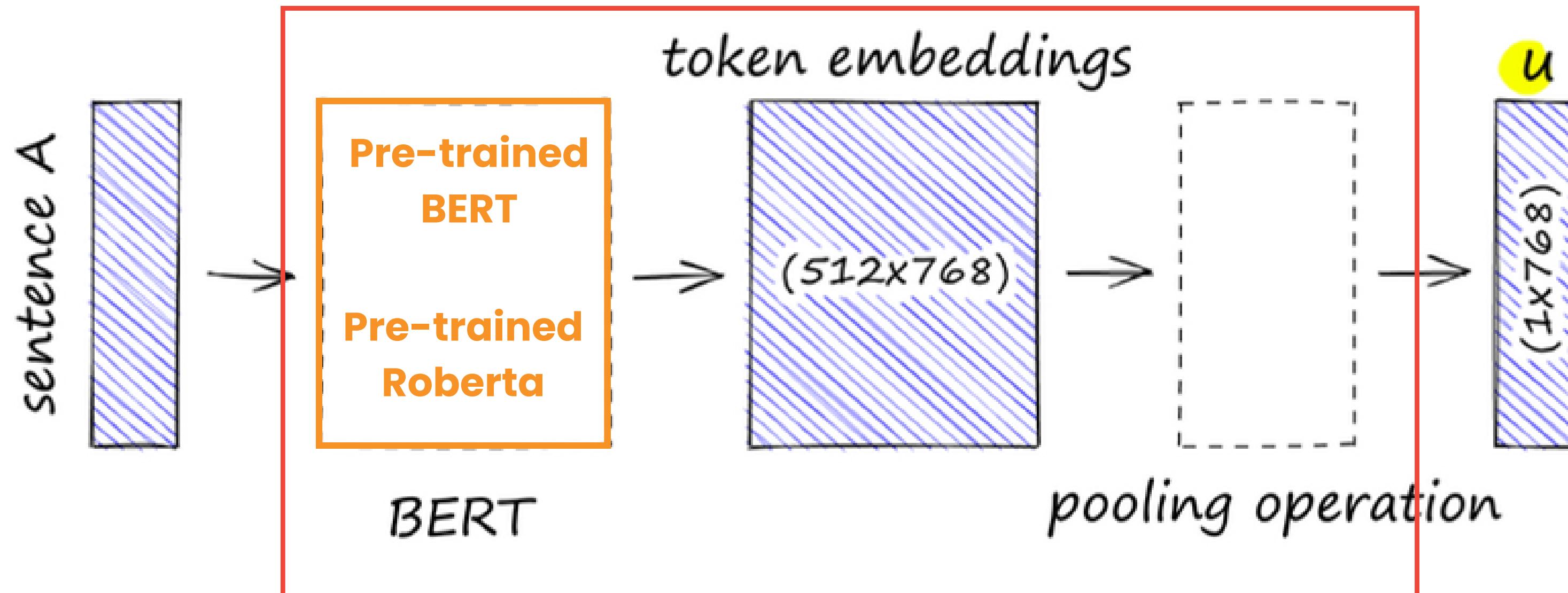
MODEL

TRAIN: SENTENCE TRANSFORMERS



MODEL

TRAIN: SENTENCE TRANSFORMERS



SENTENCE TRANSFORMERS

MODEL

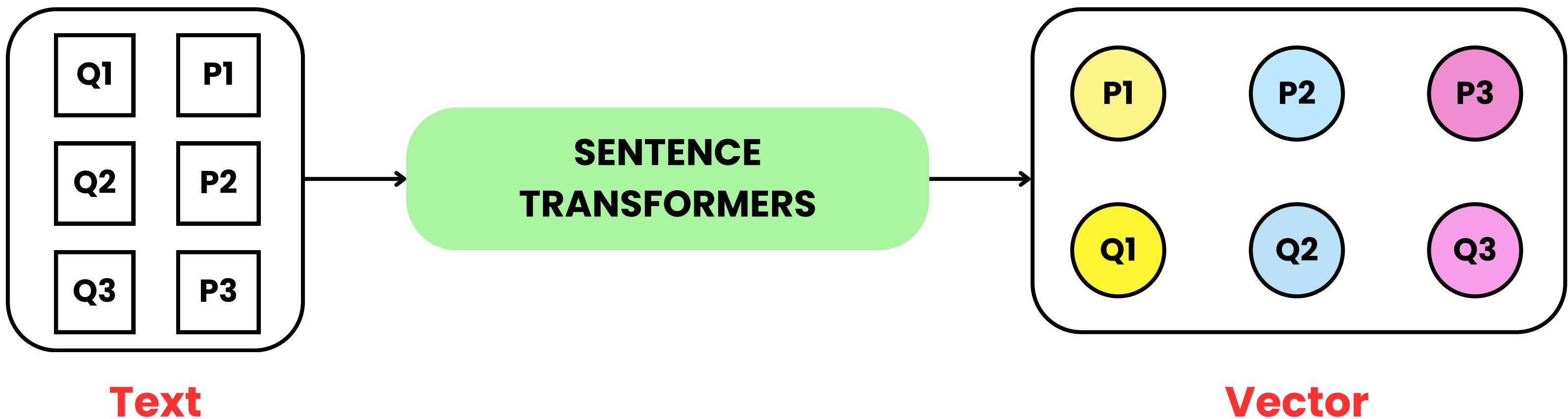
TRAIN: SENTENCE TRANSFORMERS

context string	question string
<p>1839년 바그너는 고테의 파우스트을 처음 읽고 그 내용에 마음이 끌려 이를 소재로 해서 하나의 교향곡을 쓰려는 뜻을 갖는다. 이 시기 바그너는 1838년에 빛 독촉으로 산전 수전을 다 겪은 상황이라 좌절과 실망에 가득했으며 메피스 토펠레스를 만나는 파우스트의 심경에 공감했다고 한다. 또한 파리에서 아브네크의 지휘로 파리 음악원 관현악단이 연주하는 베토벤의 교향곡 9번을 듣고 깊은 감명을 받았는데, 이것이 이듬해 1월에 파우스트의 서곡으로 쓰여진 이 작품에 조금이라도 영향을 끼쳤으리라는 것은 의심할 여지가 없다. 여기의 라단조 조성의 경우에 두 그의 전기에 적혀 있는 것</p>	<p>바그너는 고�ete의 파우스트를 읽고 무엇을 쓰고자 했는가?</p>

KorQuAD

MODEL

TRAIN: SENTENCE TRANSFORMERS

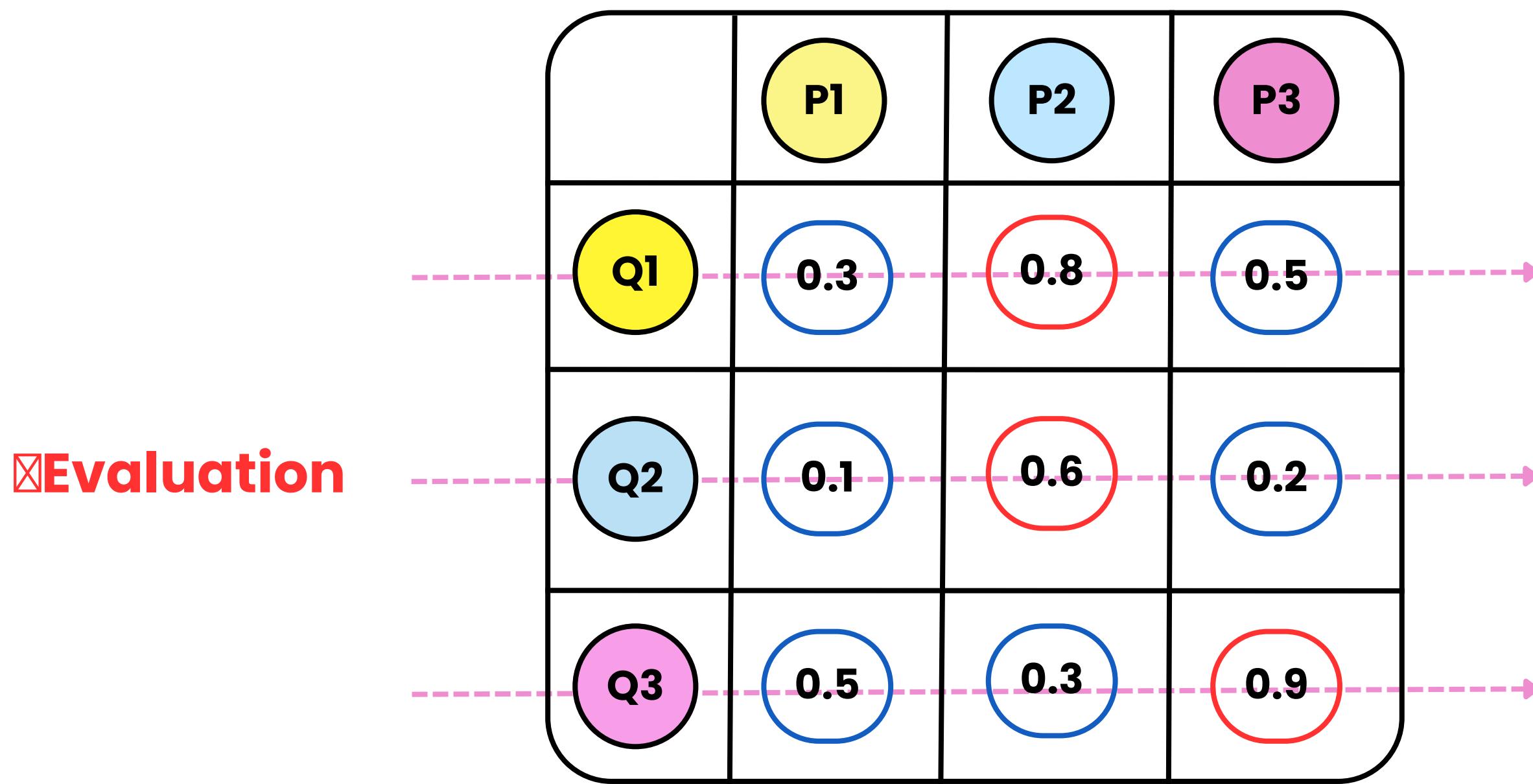


MODEL

RECAP: SENTENCE TRANSFORMERS

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|}$$

COSINE SIM



MODEL

TRAIN RESULT: SENTENCE TRANSFORMERS

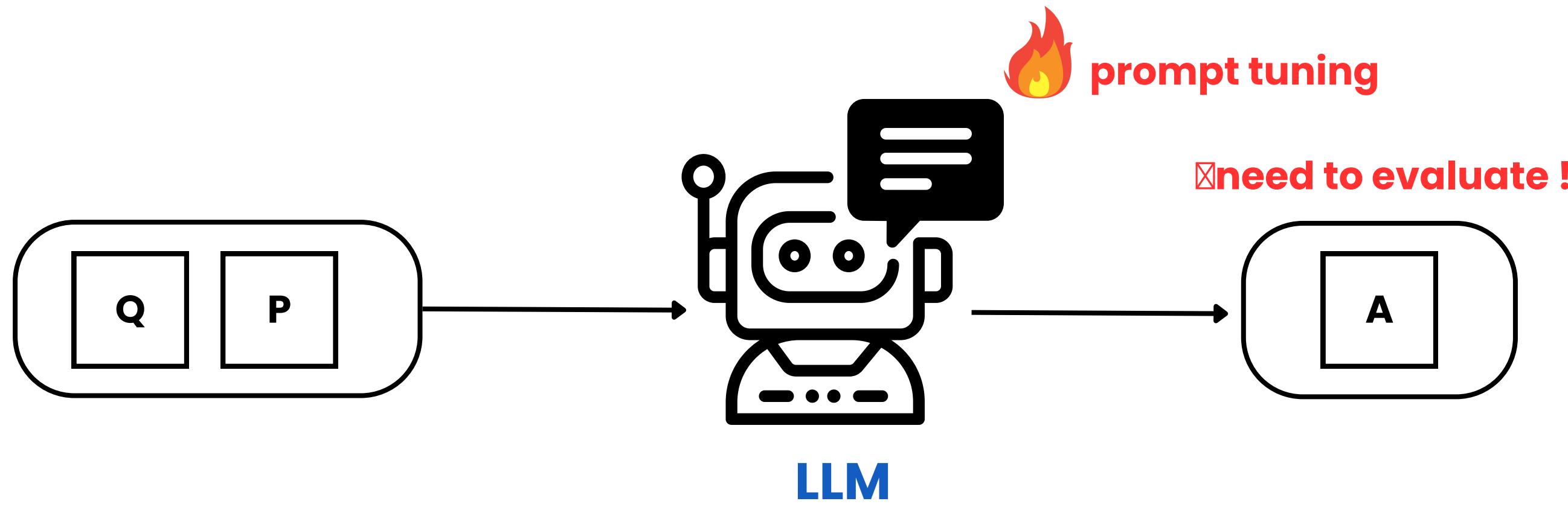
MODEL

TRAIN DATASET	BERT		RoBERTa
	Baseline (pre-trained)	0.791	0.725
	Klue-nli (pre-trained + fine-tuned)	0.873	0.869
	Ours (pre-trained + fine-tuned)	0.888	0.888

Metrics: Accuracy

MODEL

PROMPT TUNING: LLM



EXAMPLE

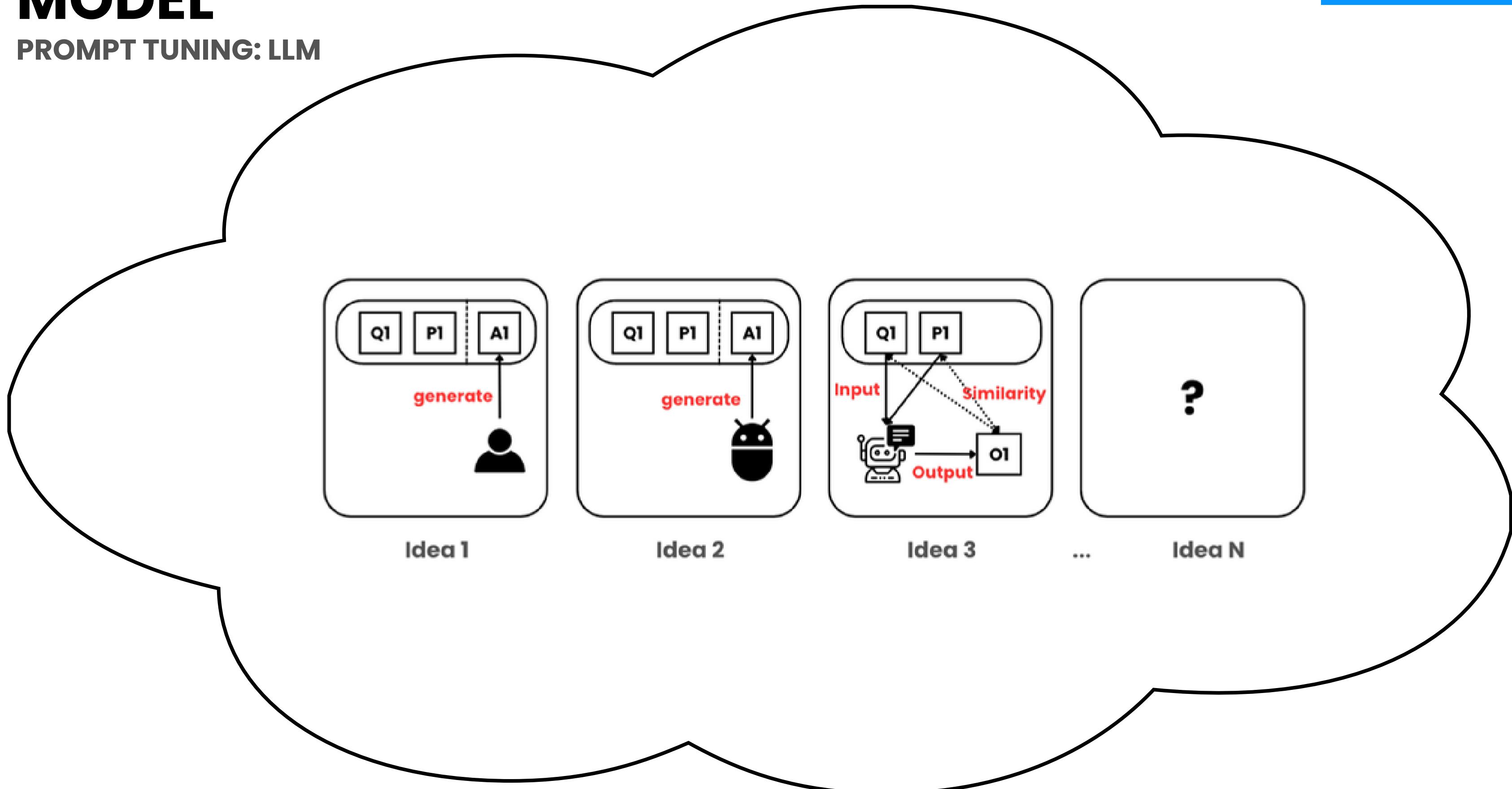
Q1: What's Sungkyunkwan University?

P1: Sungkyunkwan University is one of the
best universities in Korea.

A1: One of the best Univ in Korea.

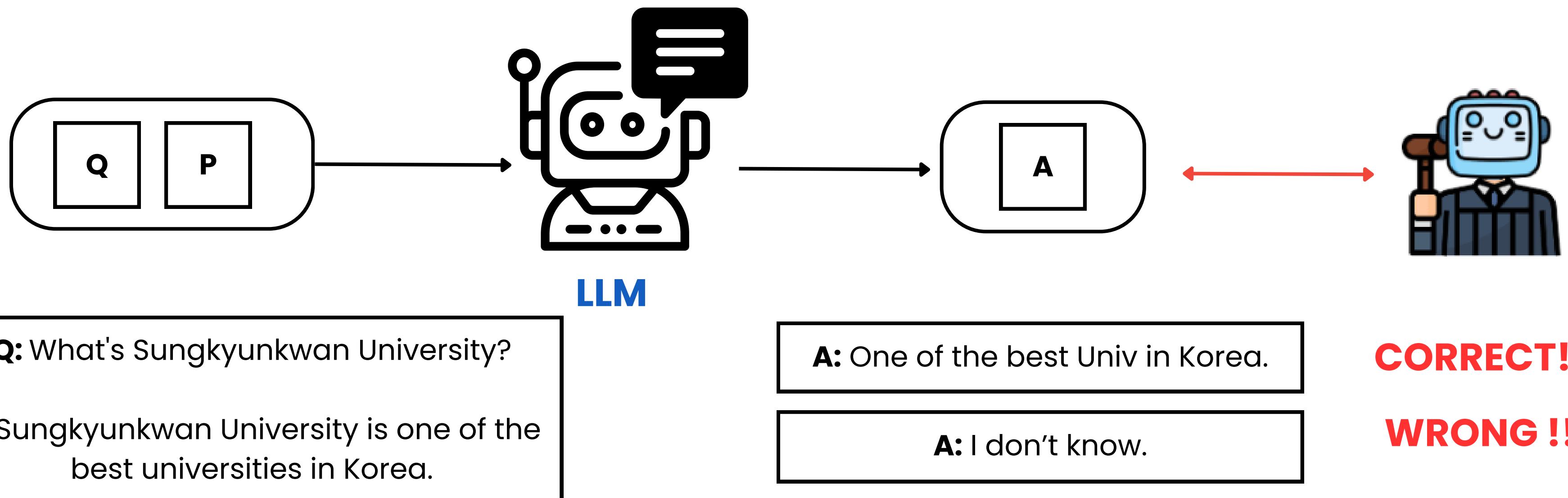
MODEL

PROMPT TUNING: LLM



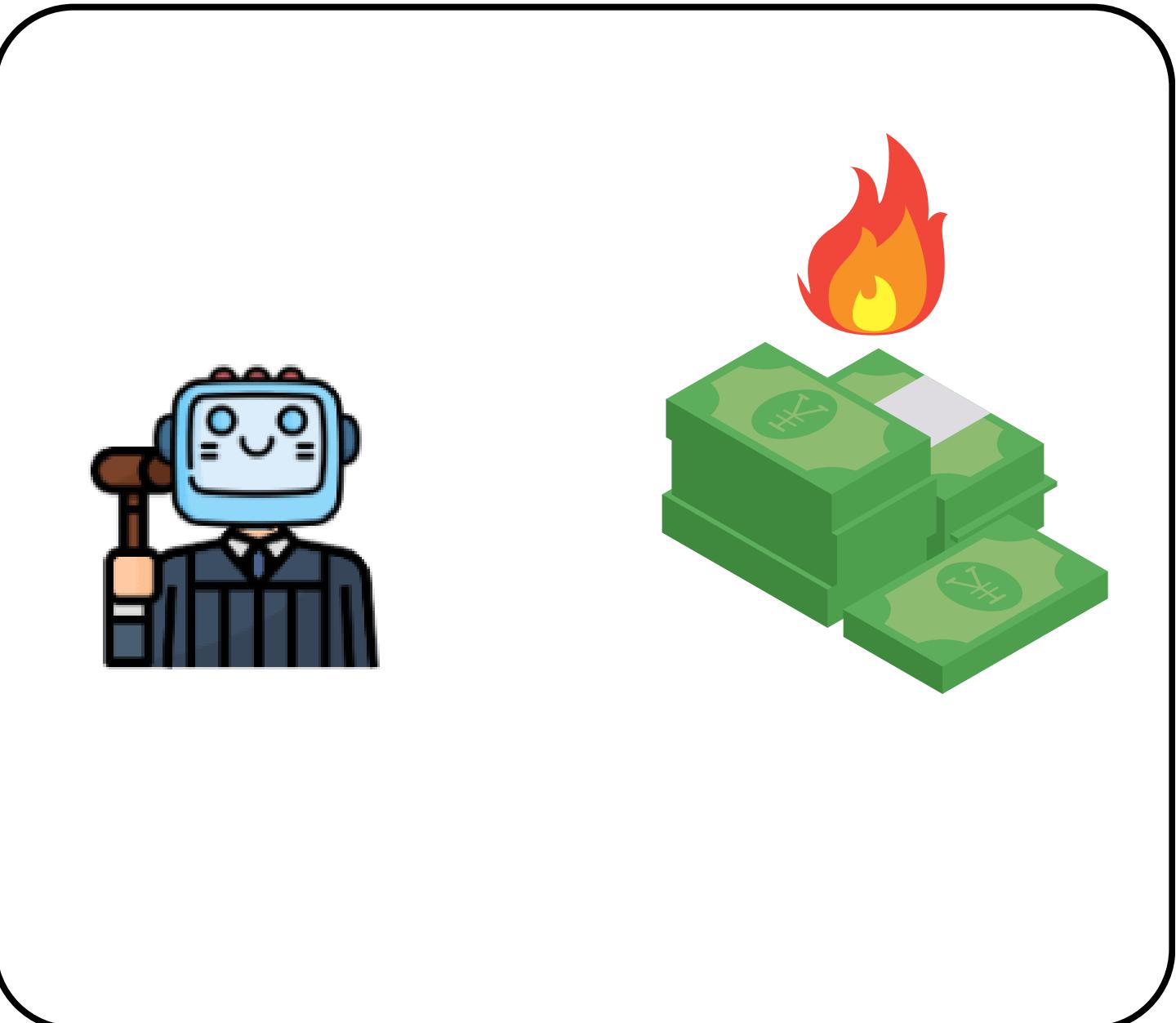
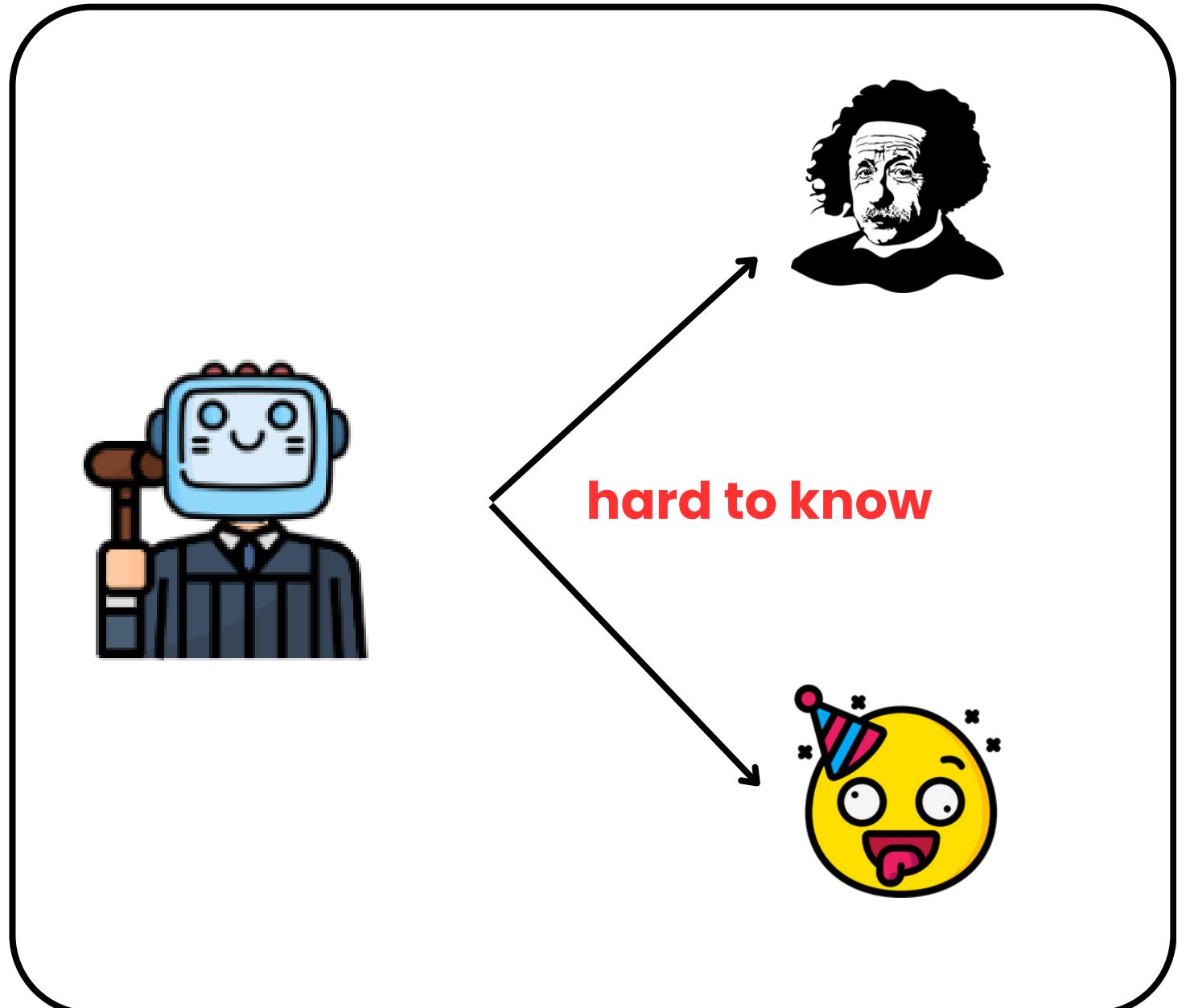
MODEL

LLM-GUIDED EVALUATION



MODEL

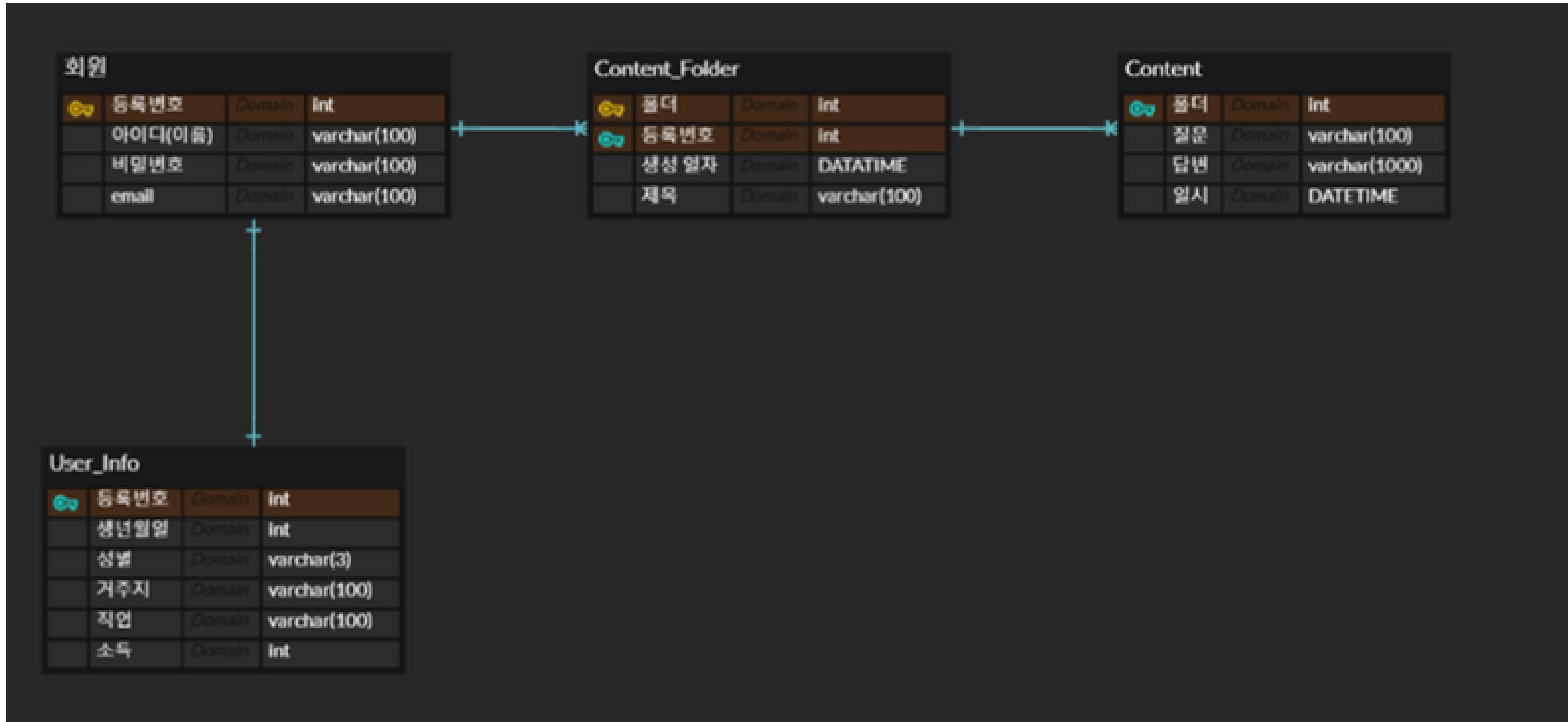
PROBLEM: LLM-GUIDED EVALUATION



BACKEND COMPONENT

- **USE FASTAPI**
 - **LOG IN, SIGN-UP , USER_INFO**
 - **FOLDER AND FOLDER CONTENT**
- **SCHEDULE : INTEGRATION AND AWS**

BACKEND DATABASE STRUCTURE OF USER



BACKEND

API SPECIFICATION

회원가입	POST	/api/user/create	회원을 등록한다.	<ul style="list-style-type: none"> • username - 사용자명 (사용자ID) • password1 - 비밀번호 • password2 - 비밀번호 확인 • email - 이메일 주소 	• 없음
로그인	POST	/api/user/login	로그인을 한다.	<ul style="list-style-type: none"> • username - 사용자명 (사용자ID) • password - 비밀번호 	<ul style="list-style-type: none"> • access_token - 액세스 토큰 • token_type - 토큰의 종류 (Bearer) • username - 사용자명 (사용자ID)
유저 정보 생성	POST	/api/userinfo/create	유저의 개인 정보를 저장한다.	<ul style="list-style-type: none"> • birth - 생년월일 • gender - 성별 • job - 직업 • region - 거주지 • money - 소득 	
유저 정보 업데이트	PUT	/api/userinfo/update	유저의 개인 정보를 업데이트한다.	<ul style="list-style-type: none"> • birth - 생년월일 • gender - 성별 • job - 직업 • region - 거주지 • money - 소득 	
유저 정보 조회	GET	/api/userinfo/detail	유저 개인 정보 조회	<ul style="list-style-type: none"> • user_id - 조회할 유저의 고유번호 	
유저 정보 삭제	DELETE	/api/userinfo/delete	유저 정보 삭제		

BACKEND

FASTAPI DOCS

POST /api/user/create User Create

POST /api/user/login Login For Access Token

POST /api/userinfo/create User Info Create

PUT /api/userinfo/update User Info Update

GET /api/userinfo/detail User Info Detail

DELETE /api/userinfo/delete User Info Delete

BACKEND

API SPECIFICATION

폴더 등록	POST	/api/folder/create	폴더를 등록한다.	• folder_name - 폴더 이름 • 없음	
폴더 조회	GET	/api/folder/list	폴더를 조회한다.		• FolderList 스키마
폴더 상세 조회	GET	/api/folder/detail/{folder_id}	폴더(folder_id)에 대한 상세 내역을 조회한다.	없음	• Folder 스키마
폴더 수정	PUT	/api/folder/update	폴더를 수정한다.	• folder_name - 폴더 이름 • folder_id- 폴더 번호	없음
폴더 삭제	DELETE	/api/folder/delete	폴더를 삭제한다.	• folder_id - 폴더 번호	없음

BACKEND

FASTAPI DOCS

GET /api/folder/list Folder List

GET /api/folder/detail/{folder_id} Folder Detail

POST /api/folder/create Folder Create

PUT /api/folder/update Folder Update

DELETE /api/folder/delete Folder Delete

BACKEND API SPECIFICATION

작업 내용	HTTP Method	URL	설명	Request Parameters	Response Schema
폴더 내용 조회	GET	/api/folder_content/list	폴더의 내용을 조회 한다.	• folder_id - 폴더 번호	FolderContentList 스키마
폴대 내용 상세 조회	GET	/api/folder_content/detail/{folder_content_id}	폴더 내용을 상세히 조회한다.	• folder_content_id - 폴더 내용 번호	FolderContent 스키마
폴대 내용 생성(질문)	POST	/api/folder_content/create	질문을 생성한다.	• folder_id • question • answer = ""	
폴더 내용 수정(답변)	PUT	/api/folder_content/update	답변을 질문에 추가 한다.	• folder_id • question • answer = "" • folder_content_id	
폴더 내용 삭제	DELETE	/api/folder_content/delete	삭제한다.	• folder_content_id	

BACKEND

FASTAPI DOCS

GET /api/folder_content/list Folder Content List

GET /api/folder_content/detail/{folder_content_id} Folder Content Detail

POST /api/folder_content/create Folder Content Create

PUT /api/folder_content/update Folder Content Update

DELETE /api/folder_content/delete Folder Content Delete

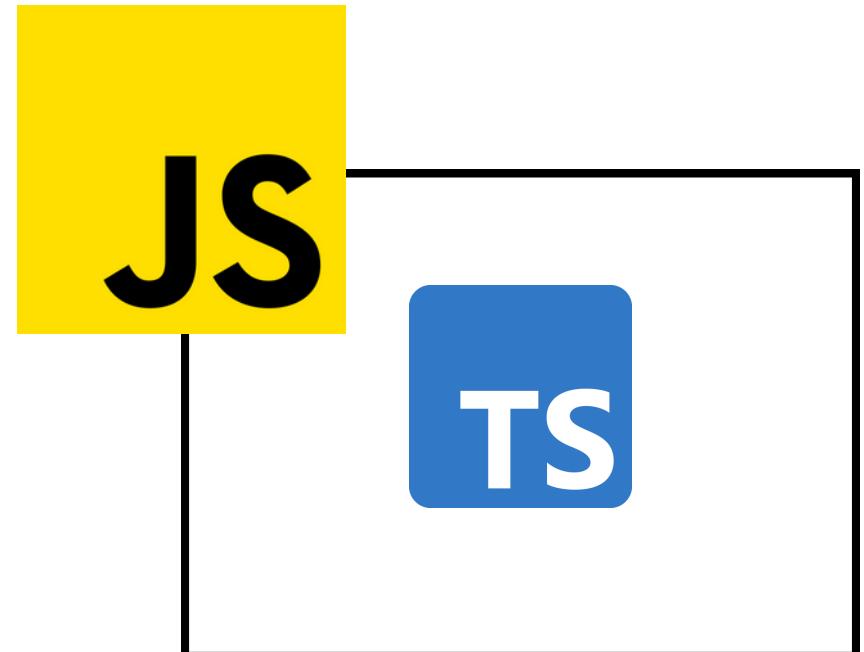
FRONT-END TOOL EXPLANATION



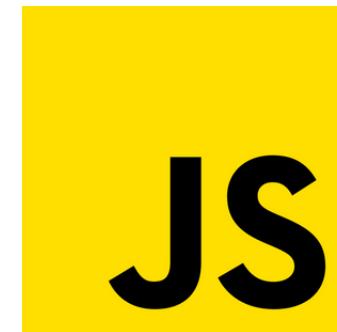
FRONT-END

TYPESCRIPT INTRO

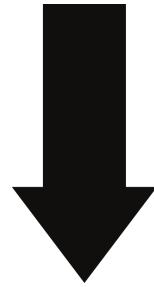
- Made by MicroSoft
- JavaScript is superset of TypeScript



FRONT-END JAVASCRIPT?

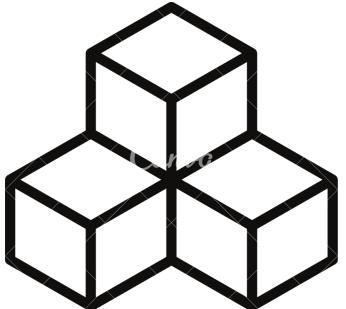


Object based Script Programming Language

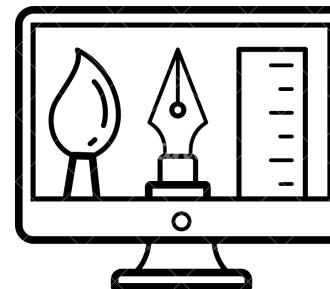


Select specific elements of HTML and add functionality (Ex : button clicks, animations)

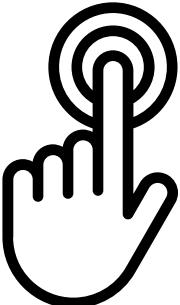
HTML : Structure of web



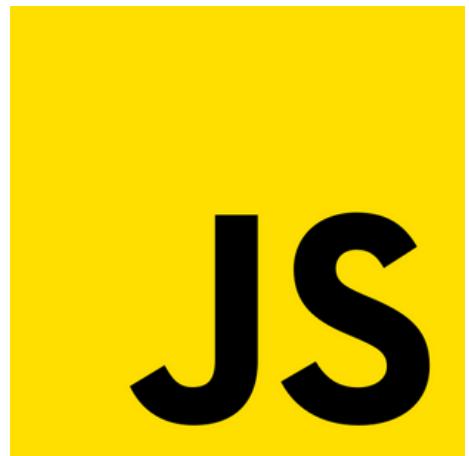
CSS : Design of web



JavaScript : Motion of web



FRONT-END FEATURE OF JS

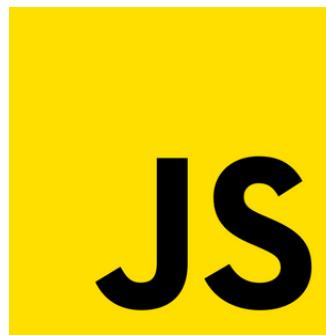


JS runs on the Web

Making it to be platform - free

Does not require strict data type checks

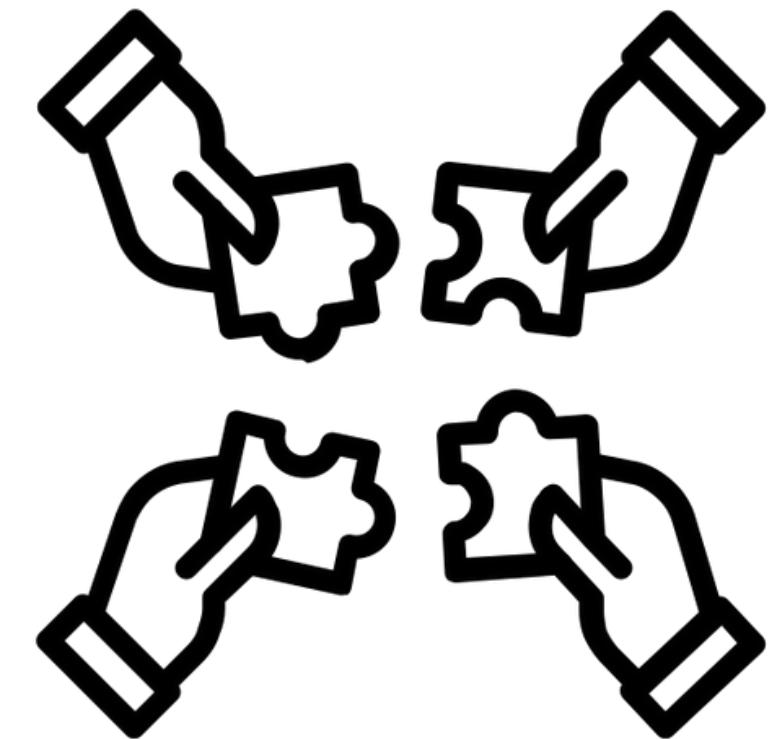
FRONT-END LIMITATION OF JS



JS runs on the Web

Making it to be platform - free

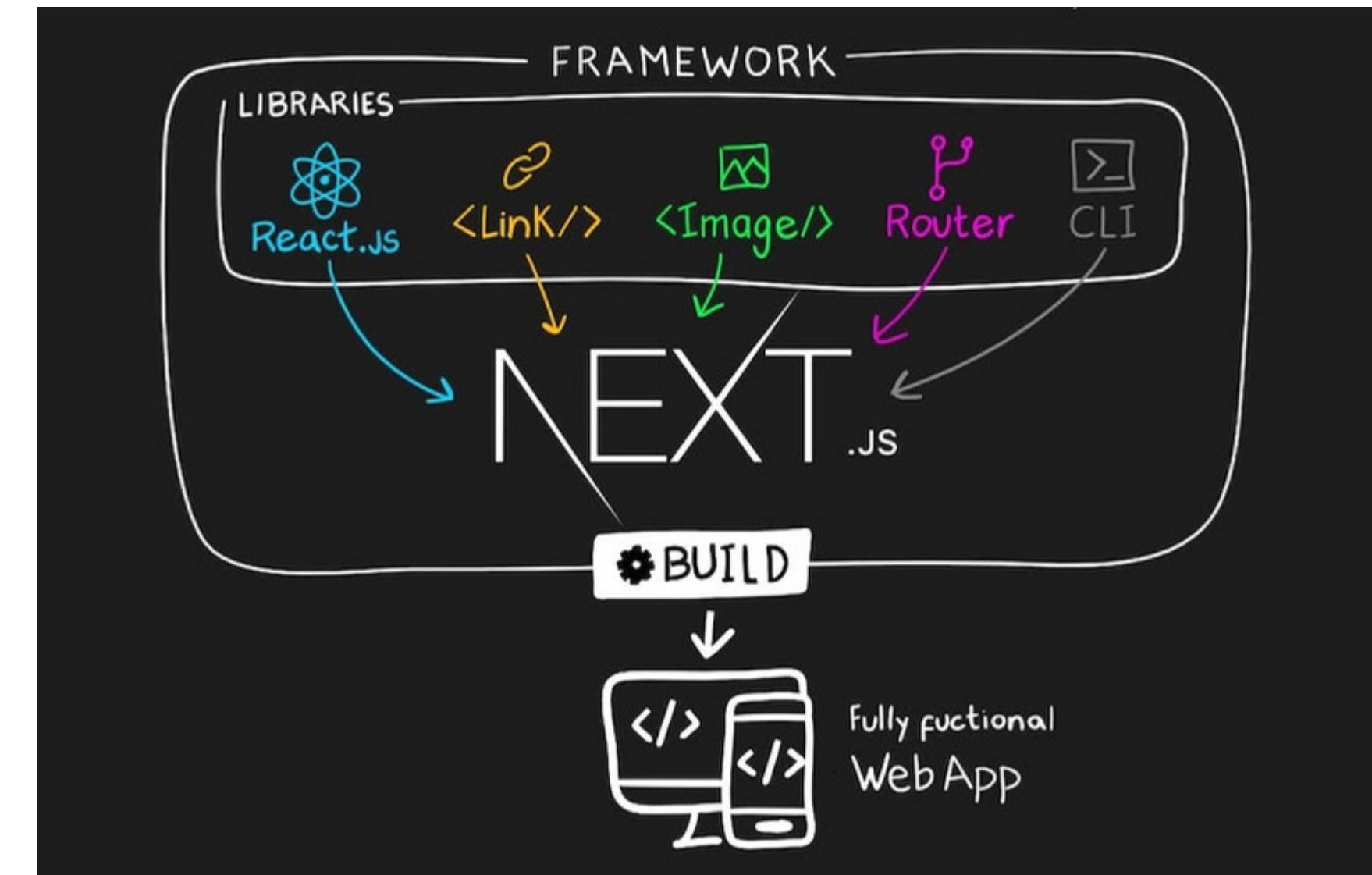
Does not require strict data type checks



$$\begin{matrix} 123 \\ abc \end{matrix} + \text{JS} = \text{TS}$$

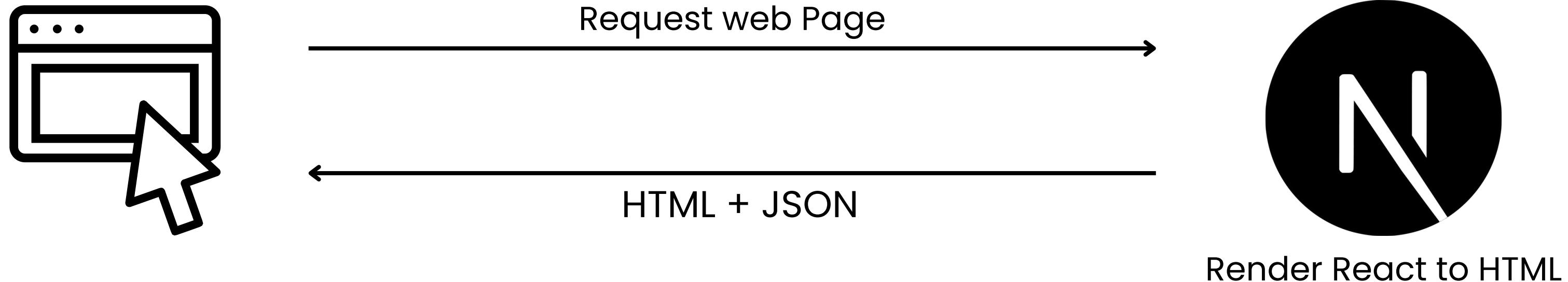
FRONT-END

NEXT.JS

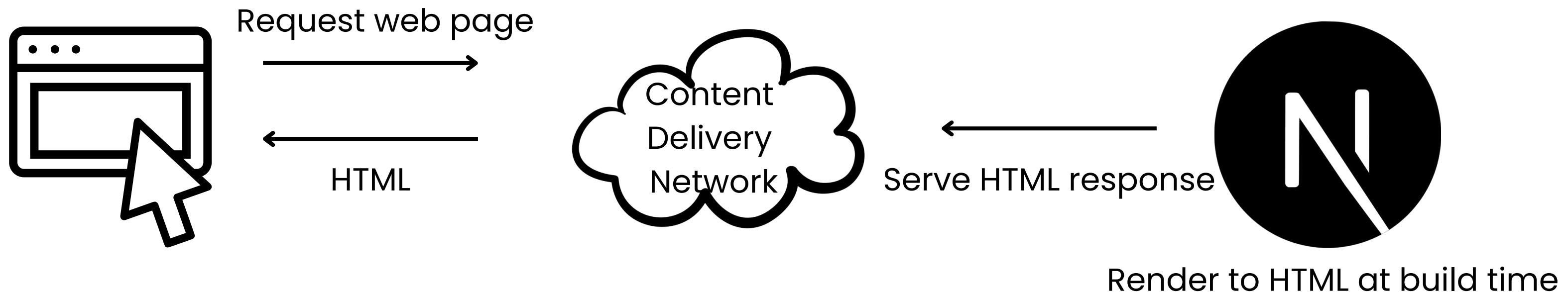


FRONT-END

NEXT.JS – SERVER SIDE RENDERING



NEXT.JS – STATIC STATIC SITE GENERATION



PROJECT SCHEDULE



QUESTIONS & ANSWERS



APPENDIX

TRAIN ARGUMENTS

Epoch: 5 (early stopping applied)

Leraning Rate: [2e-5, 5e-5, 1e-4]

Batch Size: [32, 64, 128]

Optimizer: Adam

Schedular: LinearLR